The 8-foot up and go test is the best way to assess physical function in the rheumatoid arthritis clinic

Wilkinson, Thomas; Lemmey, Andrew; Clayton, Rebecca; Jones, Jeremy; O'Brien, Thomas D.

Rheumatology Advances in Practice

DOI:
10.1093/rap/rkx017

Published: 08/01/2018

Rheumatology Advances in Practice 2018;0:1–8
doi: 10.1093/rap/rkx017

RHEUMATOLOGY ADVANCES IN PRACTICE

Original article

The 8-foot up and go test is the best way to assess physical function in the rheumatoid arthritis clinic

Thomas J. Wilkinson[1,2], Andrew B. Lemmey[1], Rebecca J. Clayton[1], Jeremy G. Jones[1,3] and Thomas D. O'Brien[4]

Abstract

Objectives. RA is characterized by poor physical function, which compromises patients' quality of life and outcome. Clinical assessment of function is usually performed using self-reported questionnaires, such as the Multi-Dimensional HAQ (MDHAQ) and the Short Form-36 (physical component) (SF36-PC). However, such subjective measures may not accurately reflect real functional status. This study aimed to determine: (i) which clinically practicable objective test best represents overall physical function; and (ii) the extent to which self-reported subjective functional measures reflect objectively assessed function.

Methods. Objective [isometric knee extensor strength, handgrip strength, sit-to-stands in 30 s, 8-foot up and go (8′UG), 50-foot walk (50′W) and estimated aerobic capacity ($\dot{V}O_2$max)] and subjective (MDHAQ and SF36-PC) measures of function were correlated with one another to determine the best overall test of functional status in 82 well-controlled RA patients (DAS28 (s.d.) = 2.8 (1.0)).

Results. In rank order of size, averaged correlations (*r*) to the other outcome measures were as follows: 8′UG: 0.650; 50′W: 0.636; isometric knee extensor strength: 0.502; handgrip strength: 0.449; sit-to-stands in 30 s: 0.432; and estimated $\dot{V}O_2$max: 0.358. The MDHAQ was weakly (0.361) and the SF36-PC moderately correlated (0.415) with objective measures.

Conclusion. Our results show that the most appropriate measure of objective physical function in RA patients is the 8′UG, followed by the 50′W. We found discordance between objectively and subjectively measured function. In clinical practice, an objective measure that is simple and quick to perform, such as the 8′UG, is advocated for assessing real functional status.

Key words: rheumatoid arthritis, physical function, objective and subjective function tests, 8-foot up and go test, health assessment questionnaire

ORIGINAL ARTICLE

Key messages

- The best measure of overall physical function is the 8′UG.
- The subjective MDHAQ and SF36-PC questionnaires are both relatively poor measures of objectively assessed physical function.
- The 8′UG and 50′W are appropriate tests for assessing functional status of RA patients in routine clinics.

[1]Rehabilitation of Musculoskeletal Disorders with Exercise Sciences (ReMeDES) group, School of Sport, Health and Exercise Sciences, Bangor University, Bangor, [2]Department of Infection, Immunity and Inflammation, University of Leicester, Leicester, [3]Peter Maddison Rheumatology Centre, Betsi Cadwaladr University Health Board, Llandudno Hospital, Llandudno and [4]Research Institute for Sport & Exercise Sciences, Liverpool John Moore's University, Liverpool, UK

Submitted 26 June 2017; revised version accepted 1 August 2017

Correspondence to: Andrew B. Lemmey, Rehabilitation of Musculoskeletal Disorders with Exercise Sciences (ReMeDES) group, School of Sport, Health and Exercise Sciences, Bangor University, Bangor, Gwynedd LL57 2PZ, UK. E-mail: a.b.lemmey@bangor.ac.uk

Introduction

RA patients typically experience adverse, inflammatory-driven changes in body composition, which result in muscle wasting and increased fat mass. This condition, termed rheumatoid cachexia (RC) [1], is a major contributor to the impaired physical function of these patients [2, 3]. Poor physical function is a strong predictor of clinical outcome in RA, and is associated with increased morbidity and mortality, reduced quality of life and higher treatment costs [4–6]. Both patients and health-care professionals identify physical function as a key outcome in the management of RA [7] and improving function as an important therapeutic target in clinical care [6, 8–10]. Hence, accurate assessment of physical function is required if RA treatment is to be optimized [10].

It was reasoned that improvements in disease management, specifically early and aggressive tight pharmaceutical control of inflammation, such as the current treat-to-target (T2T) approach [8], would attenuate RC and the associated poor physical function. However, in our recent study of 82 T2T RA patients with well-controlled disease (mean DAS28 of 2.8; 49% in clinical remission) we found that RC and markedly reduced physical function are still apparent [11]. Notably, despite these patients rating their disability as being only mild on the Multi-Dimensional HAQ (MDHAQ), strength relative to age- and sex-matched sedentary healthy controls was reduced by ~25%, and performance of objective physical function tests designed to reflect the ability to perform activities of daily living (ADL) were reduced by approximately a third. Worryingly, on average, these patients with well-controlled disease had the physical function level typical of healthy sex-matched individuals ~25 years older (based on reference values [12]).

Currently, clinical assessment of physical function in RA patients is performed using subjective, self-reported questionnaires, such as the HAQ [13] or adaptations of the HAQ (e.g. the MDHAQ [9, 14, 15]). This has been justified because of the HAQ's reliability, adaptability and ease of use and delivery, and an emphasis on outcomes that are relevant to the patient (i.e. patient-reported outcome measures) [16, 17]. Consequently, the HAQ has become the pre-eminent patient-reported outcome measure in rheumatology [15] and is recommended for the routine assessment of physical function by RA management guidelines [9, 14].

However, like most questionnaires, the HAQ suffers from various psychometric limitations; most notably, as it was designed to detect impairment among patients with greater disability than those typically seen today [17, 18], it consequently suffers from floor effects (i.e. normal scores despite physical function impairments) [15]. These effects were evident in our recent trial [11]. It is also well established that these questionnaires are influenced by age, gender, disease duration, depression, education, social class, ethnicity, personality [10, 18–21] and, especially, pain [22–25]. Notably, both the HAQ and MDHAQ are insensitive to the large functional gains (including normalization of objectively assessed function), which can result from exercise training [26–28].

Owing to the large number of external influences, it is unsurprising that self-reported function measures often show only weak-to-moderate relationships with objective performance-based tasks in RA patients [11, 19]. This weak-to-moderate association has also been demonstrated in the elderly [29, 30] and in lower back pain [31], OA [32], AS [33] and fibromyalgia patients [34]. Along with floor effects [15], it has been proposed that the discordance between the HAQ and objective measures occurs because subjective measures evaluate what patients experience when performing activities, rather than their ability to perform these activities [32]. As such, objective performance measures contribute information beyond that obtained from self-reported physical function [30] and, consequently, objective tests appear to be preferable for assessing real physical function in RA [21]. However, along with the lack of recognition and understanding of the relative advantages of objective tests over subjective questionnaires, the limitations of the clinical environment (i.e. time, space, equipment, staff and specific expertise) mitigate against the inclusion of a full battery of objective function tests in routine clinical RA management [35]. If objective function testing is to be adopted and routinely used in clinical practice, there needs to be clear guidance on which test provides the best measure of overall physical function.

Thus, the present study aimed to determine: (i) which clinically practicable objective function test best represents overall physical function; and (ii) the extent to which self-reported functional measures reflect objectively assessed function. The identification of the best overall objective physical function test in this study was exploratory and, as such, a hypothesis was not applicable. However, we did hypothesize that self-reported functional status [HAQ and SF-36 physical component (SF36-PC)] would not correlate strongly with function as indicated by objective tests.

Methods

Study population

Patients with well-controlled, stable RA were recruited from outpatient clinics of the Peter Maddison Rheumatology Centre, North Wales. For inclusion, participants had to: fulfil the ACR 2010 revised criteria for RA [36]; be aged ≥18 years; not be cognitively impaired; be free of other cachectic diseases or conditions; not be taking anabolic drugs or nutritional supplements; and not be pregnant. All patients were diagnosed after 1 January 2008 and had been exclusively treated using contemporary treatment strategies (T2T). This patient cohort was recruited for a cross-sectional study that evaluated the effect of tight control of disease activity on body composition and physical function in RA patients [11]. The study was approved by the North Wales Research Ethics Committee-West (12/WA/0323), and conducted in compliance with the Declaration of Helsinki.

Assessments and outcome measures

All outcome measures and protocols have been previously described in detail [11], so are only briefly summarized here.

Physical function

Maximal isometric knee extensor strength (IKES) was measured using an isokinetic dynamometer (Humac Cybex Norm 2004; Computer Sports Medicine Inc., MA, USA), and maximal handgrip strength (HGS) by a Grip-A dynamometer (Takei Kiki Kogyo, Japan). Objective function was also assessed by three tests specifically developed to evaluate the ability of older adults to perform ADLs [12]: sit-to-stands in 30 s (STS-30), 8-foot up and go (8′UG) and 50-foot walk (50′W). Aerobic capacity ($\dot{V}O_2$max) was estimated using the progressive submaximal Siconolfi step test [37]. This test has been validated in RA [38]. Performance of each of these tests was preceded by a submaximal practice. Our group routinely uses these tests to assess physical function in RA [11, 27, 28, 38, 39]. In order to promote inclusion into routine clinical assessments, a full step-by-step guide for delivery of each of the objective tests described here can be found in the supplementary data available at *Rheumatology Advances in Practice* online.

Subjective physical function was evaluated by the MDHAQ (scored 0–3, with higher scores denoting poorer function) as described by Pincus *et al.* [15]. The MDHAQ was used because of notable limitations of the HAQ in well-controlled RA patients; owing to its broader measurement range, this version of the HAQ has been recommended for use in patients with relatively good functional capacity, such as those in the present study [15, 17]. We also used the physical component of the SF36-PC questionnaire (scored 0–100, with lower scores denoting poorer function) [40]. The SF36-PC is the most frequently used generic scale of physical function in patients with RA [17].

Clinical measures

Disease activity was assessed by the DAS28, using CRP as the marker of systemic inflammation.

Statistical analysis

Pearson's product–moment correlation coefficient (Pearson's *r*) was used to assess the relationships between variables of interest. The correlation coefficient was measured on a scale from −1 through 0 to 1. Complete correlation between two variables is expressed by either −1 or 1, whereas complete absence of correlation is represented by 0 [41]. If a variable increases concurrently with another the correlation is positive, whereas when one variable decreases as the other increases the correlation is negative. Correlation strengths were defined according to the following criteria [41]: weak: $r = 0.200$–0.399; moderate: $r = 0.400$–0.599; and strong: $r = 0.600$–1.00.

In order to quantify the best overall measure of objective physical function, correlations (*r*) between each of the objective tests were averaged for each measure (e.g. each individual *r* value between the 8′UG and the rest of the objective tests (i.e. IKES, HGS, STS-30, 50′W and estimated $\dot{V}O_2$max) were averaged to determine a mean *r* for 8′UG. To reduce skew in the *r* distribution [42], *r* values were first transformed into normally distributed Fisher z values (z') using an established formula:

$$z' = 0.5\ln\frac{1+r}{1-r},$$

where ln is the natural logarithm function.

Following conversion to z', these values were averaged (as above) before the mean z' value was back-converted to a final *r* value ($r_{z'}$) ′using the inverse formula [42]:

$$r_{z'} = \frac{\exp^{2z'} - r}{\exp^{2z'} + r},$$

where exp is the exponential.

Subjective measures of physical function (i.e. MDHAQ and SF36-PC) were correlated with objective measures only, using the same $r_{z'}$ calculation as described above. To avoid confusion in the Results, the final summed and transformed correlation for each test (i.e. $r_{z'}$) is presented as an *r* value. Significance was set at $P < 0.05$. Data were analysed using IBM SPSS Statistics 24, and are presented as means (S.D.).

Results

Patients

Patient characteristics and mean physical function data for the 82 RA patients are shown in Table 1. Full patient demographic data are reported by Lemmey *et al.* [11].

Correlations between objective and subjective physical function measures

Table 2 displays the averaged correlations between each measure of function and the remaining objective function measures. In rank order of size, they were as follows: 8′UG: $r = 0.650$; 50′W: $r = 0.636$; IKES: $r = 0.502$; HGS: $r = 0.449$; STS-30: $r = 0.432$; and $\dot{V}O_2$max: $r = 0.358$. Although *r* values varied between males and females, 8′UG and 50′W, in that order, were the best measures of overall function (i.e. highest averaged *r* values) for both sexes, as well as for the total patient cohort. For male patients and for all patients combined, the average correlation for both 8′UG and 50′W were rated strong. Averaged correlations (male, female and combined) for IKES, HGS and STS-30 were all rated as moderate. In contrast, for both sexes and when all patients were combined, estimated $\dot{V}O_2$max was only weakly associated with the other function measures.

The MDHAQ was also only weakly correlated with the objective physical function measures ($r = 0.361$). Although the SF36-PC performed better, with a moderate averaged *r* value of 0.415 (Table 2), its average *r* was still lower than for any of the objective tests (with the exception of the submaximal $\dot{V}O_2$max step test).

Full individual correlation analysis for the total group, and for males and females, is shown in Table 3.

TABLE 1 Patient characteristics

Characteristic	Males ($n = 29$)	Females ($n = 53$)	Combined ($n = 82$)
Age, years	65.0 (7.8)	58.6 (12.9)	60.9 (11.7)
Disease duration, months	23 (18)	24 (20)	24 (19)
DAS28	2.7 (1.2)	2.9 (0.9)	2.8 (1.0)
Objective physical function			
IKES, N	474 (131)	328 (116)	380 (140)
HGS, kg	34.0 (9.6)	21.9 (5.8)	26.5 (8.8)
STS-30 (repetitions)	11.7 (4.2)	12.1 (3.3)	12.0 (3.6)
8′UG, s	7.8 (5.6)	7.2 (2.6)	7.4 (3.9)
50′W, s	11.0 (7.9)	10.6 (3.1)	10.7 (5.3)
$\dot{V}O_2$max, ml/kg/min	26.0 (7.0)	21.4 (5.2)	22.9 (6.2)
Subjective physical function			
MDHAQ (0–3)	0.64 (0.64)	0.53 (0.47)	0.57 (0.54)
SF36-PC (0–100)	42 (11)	43 (10)	43 (10)

Data are presented as means (S.D.). HGS: handgrip strength; IKES: isometric knee extensor strength; MDHAQ: multi-dimensional HAQ; SF36-PC: short-form 36 questionnaire physical component; STS-30, sit-to-stand in 30 s; 8′UG: 8-foot up and go; $\dot{V}O_2$max: estimated aerobic capacity from Siconolfi step test [data from 62 out of 82 RA patients ($n = 20$ unable to complete test: male ($n = 9$), female ($n = 11$))]; 50′W: 50-foot walk.

TABLE 2 Averaged correlations for all objective and subjective physical function measures against the objective function measures

Variable	Males ($n = 29$)	Females ($n = 53$)	Combined ($n = 82$)
IKES	0.416 (#5)	0.502 (#3)	0.502 (#3)
HGS	0.486 (#3)	0.438 (#5)	0.449 (#4)
STS-30	0.477 (#4)	0.500 (#4)	0.432 (#5)
8′UG	**0.748 (#1)**	**0.596 (#1)**	**0.650 (#1)**
50′W	0.693 (#2)	0.579 (#2)	0.636 (#2)
$\dot{V}O_2$max	0.335 (#6)	0.310 (#6)	0.358 (#6)
MDHAQ	0.428	0.369	0.361
SF36-PC	0.515	0.412	0.415

Data are presented as correlations between variables (Pearson's r). Light grey shading indicates moderate correlations ($r = 0.400$–0.599), and dark grey shading indicates strong correlations ($r = 0.600$)–1.00. (#) = rank of size for objective measures. HGS: handgrip strength; IKES: isometric knee extensor strength; MDHAQ: multi-dimensional HAQ; SF36-PC: short-form 36 questionnaire physical component; STS-30: sit-to-stand in 30 s; 8′UG: 8-foot up and go; $\dot{V}O_2$max: estimated aerobic capacity from Siconolfi step test; 50′W: 50-foot walk.

Discussion

Poor physical function is a strong predictor of clinical outcome in RA [4, 5], and accurate assessment of function is an essential part of RA management [10]. We have shown that overall objective physical function, as determined by a battery of objective function tests, is best represented by the 8′UG test followed by the 50′W. Both these tests, in male and female patients, as well as for the RA patient cohort overall, showed moderate to strong correlations with each of the other objective function tests considered. Additionally, our results show that the routine, recommended, subjective self-reported measures (e.g. MDHAQ, SF36-PC) are not representative of objectively assessed physical function.

Both the 8′UG and 50′W tests are simple to perform, require minimal equipment (i.e. markers on the floor and, in the case of the 8′UG, a standardized chair) and are easy to explain to patients. Each test is scored as time taken to complete the task (in seconds), and thus, only additionally requires a stopwatch. Both tests only take a couple of minutes to explain, and for patients to practice, and then take <1 min to perform. Unpublished observations by our group found inter-rater reliability for the 8′UG to be excellent (intraclass correlation coefficient, ICC = 0.972) and for the 50′W to be good (ICC = 0.704). Therefore, provided established protocols are followed [12] (see supplementary material 1, available at *Rheumatology Advances in Practice* online), any confounding effect of administration and scoring by different health-care professionals is small.

Although the 8′UG and 50′W were highly correlated with each other, and both performed well in comparison with the other objective function tests, we found that the 8′UG consistently demonstrated the best correlation.

TABLE 3 Correlates between physical function measures (objective and subjective) in all RA patients, male RA patients and female RA patients

	Objective physical function						Subjective physical function	
	IKES	HGS	STS-30	8′UG	50′W	$\dot{V}O_2max$	MDHAQ	SF36-PC
All RA patients								
IKES		**0.680****	**0.378***	**−0.456****	**−0.490****	**0.466****	−0.166	**0.344***
HGS	**0.680****		**0.228***	**−0.388****	**−0.363***	**0.515****	**−0.258***	**0.237***
STS-30	**0.378***	**0.228***		**−0.641****	**−0.553****	**0.289***	**−0.262***	**0.410***
8′UG	**−0.456****	**−0.388****	**−0.641****		**0.958****	**−0.285***	**0.546****	**−0.536****
50′W	**−0.490****	**−0.363***	**−0.553****	**0.958****		−0.203	**0.522****	**−0.515****
$\dot{V}O_2max$	**0.466****	**0.515****	**0.289***	**−0.285***	−0.203		−0.171	0.137
MDHAQ	−0.166	**−0.258***	**−0.262***	**0.546****	**0.522****	−0.171		**−0.627****
SF36-PC	**0.344***	**0.237***	**0.410***	**−0.536****	**−0.515****	0.137	**−0.627****	
Male RA patients								
IKES		**0.522***	0.370	**−0.422***	**−0.405***	0.354	−0.089	**0.429***
HGS	**0.522***		0.252	**−0.646****	**−0.578***	0.371	**−0.522***	**0.499***
STS-30	0.370	0.252		**−0.707****	**−0.607***	0.348	−0.169	**0.392***
8′UG	**−0.422***	**−0.646****	**−0.707****		**0.980****	−0.415	**0.644****	**−0.622****
50′W	**−0.405***	**−0.578***	**−0.607***	**0.980****		−0.173	**0.593***	**−0.590****
$\dot{V}O_2max$	0.354	0.371	0.348	−0.415	−0.173		−0.283	0.118
MDHAQ	−0.089	**−0.522***	−0.169	**0.644****	**0.593***	−0.283		**−0.682****
SF36-PC	**0.429***	**0.499***	**0.392***	**−0.622****	**−0.590****	0.118	**−0.682****	
Female RA patients								
IKES		**0.556****	**0.517****	**−0.515****	**−0.534****	**0.376***	−0.340*	**0.374***
HGS	**0.556****		**0.447****	**−0.380****	**−0.350***	**0.445****	**−0.285***	**0.195**
STS-30	**0.517****	**0.447****		**−0.620****	**−0.565****	**0.316***	**−0.331***	**0.427***
8′UG	**−0.515****	**−0.380****	**−0.620****		**0.909****	−0.217	**0.422***	**−0.504****
50′W	**−0.534****	**−0.350***	**−0.565****	**0.909****		−0.179	**0.459***	**−0.528****
$\dot{V}O_2max$	**0.376***	**0.445****	**0.316***	−0.217	−0.179		−0.162	0.201
MDHAQ	−0.340*	**−0.285***	**−0.331***	**0.422***	**0.459***	−0.162		**−0.593****
SF36-PC	**0.374***	**0.195**	**0.427***	**−0.504****	**−0.528****	0.201	**−0.593****	

Data are presented as correlations between variables (Pearson's *r*). Light grey shading indicates moderate correlations ($r = 0.400$–0.599), and dark grey shading indicates strong correlations ($r = 0.600$–1.00). $*P < 0.05$. $**P < 0.001$. Significant correlations are in bold. HGS: handgrip strength; IKES: isometric knee extensor strength; MDHAQ: multi-dimensional HAQ; SF36-PC: short-form 36 questionnaire physical component; STS-30: sit-to-stand in 30 s test; 8′UG : 8-foot up and go; $\dot{V}O_2max$: estimated aerobic capacity from Siconolfi step; 50′W: 50-foot walk.

This is likely to be because the 8′UG requires a range of physical abilities, including lower body strength, dynamic balance, walking ability, agility and gait speed [12, 43]. It also replicates an ADL patients would be required to perform several times a day (e.g. getting up from a chair to answer the telephone, go to the toilet, answer the front door, etc.). Although the 50′W may be more appropriate for some patients (e.g. those with difficulties getting out of a chair), its key limitation is the requirement for an appropriate and safe location to perform the test (i.e. a straight corridor of at least 25 feet (there and back test) free of hospital or clinic traffic). As such, especially in smaller locations, such as outpatient clinic rooms, we recommend the 8′UG for assessing objective physical function in RA patients.

In our trial, the Siconolfi step test was used to estimate cardiorespiratory fitness (i.e. $\dot{V}O_2max$). In general, this measure was poorly correlated with the other tests of physical function, which aim to reflect the ability of individuals to perform ADLs. However, as the other objective tests used were of relatively short duration (maximum of 30 s continuous physical effort for the STS-30), they are not as reliant on aerobic capacity to complete. Nevertheless, although the Siconolfi step test is not the preferred objective test for assessing disability, this test should not be overlooked. It is well established that RA patients have low aerobic capacity [44], and poor $\dot{V}O_2max$ is associated with an adverse cardiovascular profile (e.g. blood pressure, insulin resistance, high-density lipoprotein) and increased 10-year cardiovascular disease risk in RA [45]. Consequently, the submaximal Siconolfi step test, a valid and reproducible measure of $\dot{V}O_2max$ in RA [38], is a useful means of evaluating a key cardiovascular disease risk factor, namely aerobic capacity, in RA patients.

We found a discordance between subjective and objective measures of physical function. In particular, the MDHAQ was only weakly associated with objective measures. The HAQ is the recommended measure of physical function in current RA management [9, 13–16].

However, the self-reported nature of the HAQ means that it is influenced by many external factors, such as age and sex [10, 18–21] and, especially, pain [22–25]. Furthermore, the HAQ was originally intended to detect much larger functional deficits than those typically seen in the well-controlled patients of today [18] and, as such, floor effects (i.e. normal scores despite limited function) are observed [15, 17, 18, 21]. The development of the MDHAQ was claimed to reduce these floor effects by 13% compared with the original HAQ [15], making it more suitable for patients with well-controlled disease and relatively better physical function [17]. However, previous findings by our group (i.e. mild physical function score on the MDHAQ, but substantial objective functional deficits in well-treated current patients) indicate that floor effects persist in the MDHAQ [11].

Previous research into the relationship between subjective and objective measures in RA has shown large discordance between the HAQ and the observed ability of patients to perform tasks in the HAQ (e.g. tying shoelaces, doing up buttons) [19]. Objective function tests also correlate better with joint damage. Arvidson *et al.* [21] found that grip strength, walking capability and the ability to get out of a chair were strongly associated with radiographic joint damage (assessed by the modified Larsen score), whereas no correlation was found between Larsen and HAQ scores. Not surprisingly, these authors concluded that objective functional tests are preferable to HAQ for assessing physical function in RA patients.

Additionally, in the assessment of physical function, it is also important that the measure used is sufficiently sensitive to detect changes after treatment or intervention. Although self-reported physical function (e.g. the HAQ) generally shows improvement when a patient responds to pharmaceutical treatment (especially when the patient initially has uncontrolled disease activity), this response is strongly influenced by the reduction in pain that usually accompanies reduced inflammation [25] and may not necessarily reflect improvement in physical function capacity. Conversely, it is known that in patients with controlled RA, the HAQ and MDHAQ are insensitive to the large objective functional gains (including normalization of function) that result from high-intensity exercise training [26–28]. Consequently, it is important that objective, rather than subjective self-reported, measures are used to assess intervention-derived functional improvements in patients with stable RA.

As stated, the HAQ and MDHAQ have notable deficiencies, including floor effects [11, 15], considerable confounding factors [10, 18, 23, 24] and an inability to detect functional impairment in well-controlled patients [17, 18] or identify changes after exercise training [26, 28]. Nevertheless, the HAQ is recommended by international guidelines [9, 14] and the literature [16] and remains the cornerstone of functional assessment in routine RA management because of its high internal consistency, test–retest reliability, responsiveness to change in disease activity [17] and, from the clinician's perspective, limited need for training and quick ease of delivery and evaluation [15, 16]. Despite the shortcomings of subjective measures, the patient perception of physical function is an important component of medical care, as greater perceived difficulty in performing functional tasks is associated with poorer global and psychological status, dissatisfaction, reduced quality of life, loss of independence and mortality [15, 16, 30].

In comparison to the MDHAQ, we found that the SF36-PC was a marginally better indicator of overall objective physical functioning. In addition, the full SF36 has the ability to measure mental health (not measured here) and may be more relevant in assessing the economic impacts (i.e. via quality-adjusted life-years) of treatment [46], which may be particularly pertinent to RA, where expensive biologic treatment is frequently used. Nonetheless, although the SF36-PC may be a more accurate depiction of physical function than the HAQ in RA, it is unlikely to be widely adopted into routine rheumatology practice for the following reasons: (i) it is a generic instrument, whereas the HAQ and its derivatives are RA specific [17]; (ii) there are expensive annual licence fees for software and the questionnaire itself [47]; and (iii) relative to the HAQ, evaluating the SF36-PC is arduous and requires specialized scoring software.

The lack of agreement between objectively and subjectively assessed function reported in RA [19, 21], the elderly [29, 30] and other rheumatic conditions, including back pain, OA, AS and FM [33], may arise from inaccurate reporting, limited response criteria and/or measurement error or, alternatively, because performance-based measures may not assess all aspects that influence a participant's subjective functional assessment of that task [30]. Accordingly, because objective and subjective measures provide information on different aspects of physical function [21, 32], they may be complementary and, in combination, should provide a more comprehensive understanding of a person's functional status [30]. However, owing to the numerous deficiencies of subjective measures, an objective measure should routinely be assessed.

A limitation of the present study is the largely homogeneous cohort assessed. Patients were generally well controlled [mean DAS28-CRP (S.D.) = 2.8 (1.0), 49% in remission] and, in this regard, our finding that the 8′UG is the most appropriate function test may only be applicable to patients with stable and generally low disease activity. Although the 8′UG remains a potentially feasible and practical test for patients with high disease activity to perform, additional research should confirm whether this test is also the best overall measure of physical function in this group of patients. Future research could also investigate relevant cut-off points for objective physical function performance in RA patients. Previous work in the elderly [30, 48] and other clinical populations, such as Parkinson's disease [43] and diabetic peripheral neuropathy [49], has shown that 8′UG cutoffs can accurately predict functional mobility skills (e.g. getting into a shower) and fall risk.

Conclusion

Our results show that the 8′UG, followed by the 50′W, demonstrate the strongest correlations with other measures of physical function. Consequently, the 8′UG appears to be the most appropriate objective assessment to assess overall physical function of patients with RA. We found a discordance between function measured by objective tests and that measured subjectively by the MDHAQ and SF36-PC questionnaires. Consequently, in clinical practice, the addition of an objective measure, such as the 8′UG, which is simple and quick to perform, is advocated for determining a patient's real physical function, and may provide a more accurate and comprehensive picture of a patient's functional status.

Funding: This research was not funded by a specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Disclosure statement: The authors declare no conflict of interest.

Supplementary data

Supplementary data are available at *Rheumatology Advances in Practice* online.

References

1 Roubenoff R, Roubenoff RA, Cannon JG et al. Rheumatoid cachexia: cytokine-driven hypermetabolism accompanying reduced body cell mass in chronic inflammation. *J Clin Invest* 1994;93:2379–86.

2 Giles JT, Bartlett SJ, Andersen RE et al. Association of body composition with disability in rheumatoid arthritis: impact of appendicular fat and lean tissue mass. *Arthritis Rheum* 2008;59:1407–15.

3 Kramer HR, Fontaine KR, Bathon JM, Giles JT. Muscle density in rheumatoid arthritis: associations with disease features and functional outcomes. *Arthritis Rheum* 2012; 64:2438–50.

4 Gabriel SE, Crowson CS, Kremers HM et al. Survival in rheumatoid arthritis: a population-based analysis of trends over 40 years. *Arthritis Rheum* 2003;48:54–8.

5 Pincus T, Castrejón I. Evidence that the strategy is more important than the agent to treat rheumatoid arthritis. *Bull Hosp Jt Dis* 2013;71:S33–40.

6 Aletaha D, Smolen J, Ward MM. Measuring function in rheumatoid arthritis: identifying reversible and irreversible components. *Arthritis Rheum* 2006;54:2784–92.

7 Bartlett SJ, Hewlett S, Bingham CO III et al. Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. *Ann Rheum Dis* 2012;71:1855–60.

8 Smolen JS, Breedveld FC, Burmester GR et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;75:3–15.

9 Singh J A, Furst D E, Bharat A et al. 2012 Update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. *Arthritis Care Res* 2012;64:625–39.

10 Liegl G, Gandek B, Fischer HF et al. Varying the item format improved the range of measurement in patient-reported outcome measures assessing physical function. *Arthritis Res Ther* 2017;19:66.

11 Lemmey AB, Wilkinson TJ, Clayton RJ et al. Tight control of disease activity fails to improve body composition or physical function in rheumatoid arthritis patients. *Rheumatology* 2016;55:1736–45.

12 Rikli RE, Jones CJ. Development and validation of criterion-referenced clinically relevant fitness standards for maintaining physical independence in later years. *Gerontologist* 2013;53:255–67.

13 Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23: 137–45.

14 Deighton C, O'Mahony R, Tosh J et al. Management of rheumatoid arthritis: summary of NICE guidance. *BMJ* 2009;338:b702.

15 Pincus T, Sokka T, Kautiainen H. Further development of a physical function scale on a MDHAQ [corrected] for standard care of patients with rheumatic diseases. *J Rheumatol* 2005;32:1432–39.

16 Bruce B, Fries JF. The health assessment questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23:S14–8.

17 Voshaar MA, ten Klooster PM, Taal E, van de Laar MA. Measurement properties of physical function scales validated for use in patients with rheumatoid arthritis: a systematic review of the literature. *Health Qual Life Outcomes* 2011;9:99.

18 Wahl E, Gross A, Chernitskiy V et al. Validity and responsiveness of a 10-item patient-reported measure of physical function, PROMIS PF-10a, in a Rheumatoid Arthritis Clinic Population. *Arthritis Care Res* 2017;69:338–46.

19 Van Den Ende C, Hazes JM, Le Cessie S, Breedveld FC, Dijkmans BA. Discordance between objective and subjective assessment of functional ability of patients with rheumatoid arthritis. *Br J Rheumatol* 1995;34:951–55.

20 Daltroy LH, Larson MG, Eaton HM, Phillips CB, Liang MH. Discrepancies between self-reported and observed physical function in the elderly: the influence of response shift and other factors. *Soc Sci Med* 1999;48:1549–61.

21 Arvidson N G, Larsson A, Larsen A. Simple function tests, but not the modified HAQ, correlate with radiological joint damage in rheumatoid arthritis. *Scand J Rheumatol* 2002;31:146–50.

22 Malm K, Bergman S, Andersson M, Bremander A; BARFOT Study Group. Predictors of severe self-reported disability in RA in a long-term follow-up study. *Disabil Rehabil* 2015;37:686–91.

23 Sokka T, Kankainen A, Hannonen P. Scores for functional disability in patients with rheumatoid arthritis are

correlated at higher levels with pain scores than with radiographic scores. *Arthritis Rheum* 2000;43:386–9.

24 Wolfe F. A reappraisal of HAQ disability in rheumatoid arthritis. *Arthritis Rheum* 2000;43:2751–61.

25 Scott D, Strand V. The effects of disease-modifying anti-rheumatic drugs on the Health Assessment Questionnaire score. Lessons from the leflunomide clinical trials database. *Rheumatology* 2002;41:899–909.

26 Van den Ende C, Breedveld FC, Dijkmans BA, Hazes JM. The limited value of the Health Assessment Questionnaire as an outcome measure in short term exercise trials. *J Rheumatol* 1997;24:1972–77.

27 Marcora SM, Lemmey AB, Maddison PJ. Can progressive resistance training reverse cachexia in patients with rheumatoid arthritis? Results of a pilot study. *J Rheumatol* 2005;32:1031–9.

28 Lemmey AB, Marcora SM, Chester K et al. Effects of high-intensity resistance training in patients with rheumatoid arthritis: a randomized controlled trial. *Arthritis Care Res* 2009;61:1726–34.

29 Hoeymans N, Feskens EJ, van den Bos GA, Kromhout D. Measuring functional status: cross-sectional and longitudinal associations between performance and self-report (Zutphen Elderly Study 1990–1993). *J Clin Epidemiol* 1996;49:1103–10.

30 Guralnik JM, Simonsick EM, Ferrucci L et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994;49:M85–94.

31 Agwubike E, Ezeukwu A. Comparison of subjective and objective physical functions in patients with chronic low back pain. *Online J Health Allied Scs* 2011;10(2).

32 Terwee CB, Mokkink LB, Steultjens P, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Rheumatology* 2006;45:890–902.

33 van Weely SF, van Denderen JC, Steultjens MP et al. Moving instead of asking? Performance-based tests and BASFI-questionnaire measure different aspects of physical function in ankylosing spondylitis. *Arthritis Res Ther* 2012;14:R52.

34 Hidding A, van Santen M, De Klerk E et al. Comparison between self-report measures and clinical observations of functional disability in ankylosing spondylitis, rheumatoid arthritis and fibromyalgia. *J Rheumatol* 1994;21:818–23.

35 Lemmey AB. Rheumatoid cachexia: the undiagnosed, untreated key to restoring physical function in rheumatoid arthritis patients? *Rheumatology* 2016;55:1149–50.

36 Aletaha D, Neogi T, Silman AJ et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.

37 Siconolfi SF, Garber CE, Lasater TM, Carleton RA. A simple, valid step test for estimating maximal oxygen uptake in epidemiologic studies. *Am J Epidemiol* 1985;121:382–90.

38 Cooney JK, Moore JP, Ahmad YA et al. A simple step test to estimate cardio-respiratory fitness levels of rheumatoid arthritis patients in a clinical setting. *Int J Rheumatol* 2013;2013:174541.

39 Wilkinson TJ, Lemmey AB, Jones JG et al. Can creatine supplementation improve body composition and objective physical function in rheumatoid arthritis patients? A randomized controlled trial. *Arthritis Care Res* 2016;68:729–37.

40 Brazier JE, Harper R, Jones NM et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305:160–64.

41 Campbell M J, Swinscow T D V. Correlation and regression. In: *Statistics at square one*, 9th edn. London: BMJ Publishing Group, 2009, 119–33.

42 Corey DM, Dunlap WP, Burke MJ. Averaging correlations: expected values and bias in combined Pearson *r*s and Fisher's *z* transformations. *J Gen Psychol* 1998;125:245–61.

43 Nocera JR, Stegemöller EL, Malaty IA et al. Using the timed Up & Go test in a clinical setting to predict falling in Parkinson's disease. *Arch Phys Med Rehabil* 2013;94:1300–05.

44 Munsterman T, Takken T, Wittink H. Are persons with rheumatoid arthritis deconditioned? A review of physical activity and aerobic capacity. *BMC Musculoskelet Disord* 2012;13:202.

45 Metsios GS, Koutedakis Y, Veldhuijzen van Zanten JJ et al. Cardiorespiratory fitness levels and their association with cardiovascular profile in patients with rheumatoid arthritis: a cross-sectional study. *Rheumatology* 2015;54:2215–20.

46 Scott DL, Khoshaba B, Choy EH, Kingsley GH. Limited correlation between the Health Assessment Questionnaire (HAQ) and EuroQol in rheumatoid arthritis: questionable validity of deriving quality adjusted life years from HAQ. *Ann Rheum Dis* 2007;66:1534–7.

47 Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: scoping review. *SAGE Open Med* 2016;4:2050312116671725.

48 Shumway-Cook A, Baldwin M, Polissar NL, Gruber W. Predicting the probability for falls in community-dwelling older adults. *Phys Ther* 1997;77:812–9.

49 Jernigan SD, Pohl PS, Mahnken JD, Kluding PM. Diagnostic accuracy of fall risk assessment tools in people with diabetic peripheral neuropathy. *Phys Ther* 2012;92:1461–70.